

# A Theory for Optical flow-based Transport on Image Manifolds

Sriram Nagaraj, Aswin C. Sankaranarayanan, Richard G. Baraniuk

*Rice University*

---

## Abstract

An image articulation manifold (IAM) is the collection of images formed when an object is articulated in front of a camera. IAMs arise in a variety of image processing and computer vision applications, where they provide a natural low-dimensional embedding of the collection of high-dimensional images. To date IAMs have been studied as embedded submanifolds of Euclidean spaces. Unfortunately, their promise has not been realized in practice, because real world imagery typically contains sharp edges that render an IAM non-differentiable and hence non-isometric to the low-dimensional parameter space under the Euclidean metric. As a result, the standard tools from differential geometry, in particular using linear tangent spaces to transport along the IAM, have limited utility. In this paper, we explore a nonlinear transport operator for IAMs based on the *optical flow* between images and develop new analytical tools reminiscent of those from differential geometry using the idea of *optical flow manifolds* (OFMs). We define a new metric for IAMs that satisfies certain local isometry conditions, and we show how to use this metric to develop a new tools such as flow fields on IAMs, parallel flow fields, parallel transport, as well as a intuitive notion of curvature. The space of optical flow fields along a path of constant curvature has a natural multi-scale structure via a monoid structure on the space of all flow fields along a path. We also develop lower bounds on approximation errors while approximating non-parallel flow fields by parallel flow fields.

**Keywords:** Image articulation manifolds, Transport operators, Optical flow

---

## 1. Introduction

### 1.1. Image articulation manifolds

Many problems in image processing and computer vision involve image ensembles that are generated by varying a small set of imaging parameters such as pose, lighting, view angle, etc. of a fixed three-dimensional (3D) scene. As the parameters vary, the images can be modeled as a lying on a (typically nonlinear) manifold called an *image articulation manifold (IAM)* [1, 2, 3, 4]. Each point on an IAM is an image at a particular parameter value. Over the past decade, there has been significant work [3, 5, 6, 7] in learning and processing the underlying geometric structures associated with image ensembles. For instance, tasks such as recognition, classification, and image synthesis can be interpreted as navigation along a particular IAM.

More specifically, we define an image articulation manifold (IAM) as the set of images  $M$  formed by the action of an imaging map  $i$  on a space of articulations  $\Theta$ , i.e.,  $M = \{i_\theta = i(\theta) : \theta \in \Theta\}$ . This imaging process can be decomposed into two steps: first the action of the articulation on a 3D object or scene and then the subsequent imaging of the articulated object/scene (see [8] for a detailed discussion of image formation).<sup>1</sup>

### 1.2. IAM Challenges

In spite of much progress, there are fundamental challenges to successfully applying manifold-processing tools to generic image data, in particular IAMs.

First, it has been shown that IAMs containing images with sharp edges are *non-differentiable* [1]. Specifically, given parameters  $\theta_1$  and  $\theta_2$  with corresponding images  $I_1$  and  $I_2$ , it follows that the  $L^2$  distance  $\|I_1 - I_2\|_{L^2}$  between images  $I_1$  and  $I_2$  is a nonlinear function  $\eta(\|\theta_1 - \theta_2\|)$  that is asymptotically equivalent to  $(\|\theta_1 - \theta_2\|)^{\frac{1}{2}}$ . Indeed,  $\frac{\|I_1 - I_2\|_{L^2}}{\|\theta_1 - \theta_2\|} \geq c\|\theta_1 - \theta_2\|^{-\frac{1}{2}}$ , and this non-Lipschitz relation

---

<sup>1</sup>Before we proceed further, it is worth discussing certain degeneracies in the imaging process that we wish to avoid in this paper for analytical reasons. In particular, we want to avoid cases where the set  $M$  is not expressive of the full range of articulations. As an example, consider a uniformly colored 3D sphere  $\mathcal{O}$  undergoing rotation about a fixed axis. Here, the parameter space  $\Theta$  is the unit circle  $S^1$ . Being uniformly colored, the sphere does not change appearance under rotation, and the IAM degenerates to a single point  $\{I_{\mathcal{O}}\}$ . Were the sphere richly textured, we would obtain new views of the sphere for each rotation so that the IAM is isomorphic to  $S^1$ . For the remainder of the paper, we will consider IAMs without degeneracies by assuming that the imaging map  $i$  is a re-parametrization of  $\Theta$ ; i.e., we will assume that the IAM is homeomorphic to the corresponding parameter space.

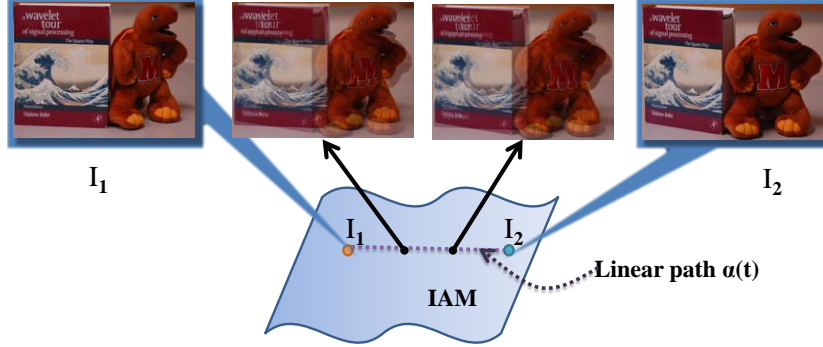


Figure 1: Non-differentiability of IAMs renders (locally) linear transport inaccurate. Consider an image interpolation task. Given images  $I_1$  and  $I_2$ , the line connecting them  $\alpha(t) = tI_1 + (1-t)I_2$  is blurred and thus a poor approximation to the geodesic connecting  $I_1$  and  $I_2$ .

indicates that the corresponding IAM is non-differentiable. Non-differentiability suggests that local linear approximations, such as those suggested by differential geometry, are invariably inaccurate on IAMs. To illustrate this, consider a stylized example of image interpolation (see Fig. 1). Given images  $I_1, I_2 \in M$ , consider the affine path  $\alpha(t) = tI_1 + (1-t)I_2$  with  $t \in [0, 1]$ ; for a smooth manifold ( $> \mathcal{C}^2$ ), this line would be a close approximation to the actual manifold, especially over small neighborhoods. However, non-differentiability implies that a first-order approximation to the manifold is inaccurate even over a small neighborhood. This is illustrated in Fig. 1. To alleviate the non-differentiability problem, Wakin et al. [2] have proposed a multiscale smoothing procedure that regularizes each point of the IAM by a set of multiscale Gaussian smoothing filters that render the IAM smooth across the various scales. This smoothing procedure then enables the definition of linear tangent spaces on which one can perform standard linear methods of analysis. However, this is unsatisfactory, since Gaussian smoothing is inherently lossy — leading to loss of high-frequency information in the images. Further, defining tangent vectors as the limit of a multiscale procedure is inherently complex and moreover not possible for practical scenarios where we have only samples from the IAM.

Second, conventional manifold models lack a meaningful metric between points on the IAM, especially when the sampling of the manifold is sparse. Consider the simple *translation manifold*  $M_T$  generated by imaging a black disk of radius  $R$  translating on an infinite white background (see Fig. 2). Let  $I_1$  be the image of the disk with center  $c_1$  and  $I_2$  the image of the disk with center  $c_2$ . It

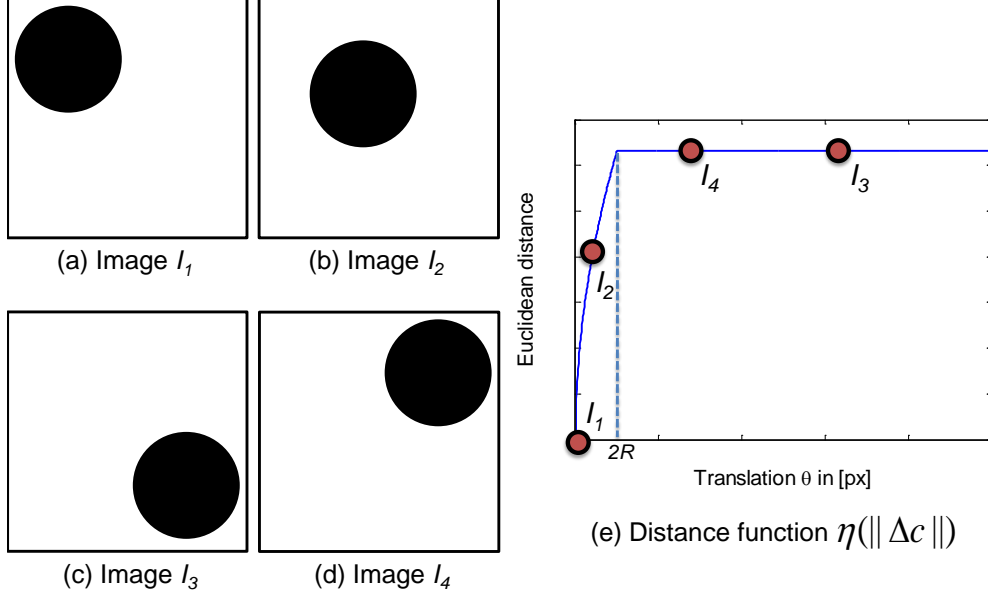


Figure 2: Euclidean distance between images on an IAM can be meaningless. (a-d) Images of a translating disk over a white background. (e) The  $L^2$  distance between two images depends on the amount of overlap between the disks. However, the overlap is zero when the distance between centers exceed  $2R$ ; in the figure above,  $d(I_1, I_3) = d(I_1, I_4)$ . Hence, the distance between two images can be written as  $\eta(\min(2R, \|c_1 - c_2\|_2))$  where  $\eta(\cdot)$  is a monotonically increasing function. Conventional manifold learning techniques, which rely heavily on meaningful local distances, do not work unless the sampling on the manifold is dense.

then follows that

$$\|I_1 - I_2\|_{L^2} = \eta(\min(2R, \|c_1 - c_2\|)) = \begin{cases} \eta(\|c_1 - c_2\|), & \|c_1 - c_2\| < 2R \\ \eta(2R), & \|c_1 - c_2\| \geq 2R \end{cases}$$

where  $\eta(\cdot)$  is a nonlinear function. Thus, when the disk centers are separated by a distance greater than the diameter of the disk, the metric is completely uninformative (see Fig. 2). This suggests that, unless the sampling of images from the IAM is sufficiently dense, organization of the images using a construct such as a  $k$ -nearest neighbor graph is meaningless.  $k$ -nearest neighbor graphs are at the heart of traditional manifold learning techniques such as LLE [6], ISOMAP [5], and diffusion maps [3]; as a consequence, such techniques are doomed to fail unless the sampling is sufficiently dense.

These shortcomings are exacerbated for images with rich textures; this rules out a consistent analysis of IAMs based on classical differential geometry that

relies on the smoothness and metric properties of the manifold. A critical missing link in the analysis of IAMs is a systematic theory that successfully handles the above issues and permits new mathematical tools for IAMs, including analogues of traditional notions such as curvature, vector fields, parallel transport, etc. [9]. Clearly, once such analytic tools are available, we can greatly expand the scope of applications and pave the way for efficient algorithms specific to IAMs.

### 1.3. Transport operators

Recently, a new class of methods for handling image ensembles has been developed based on the idea of transport operators [4, 10, 11, 12, 13]. A transport operator on an IAM is a (typically nonlinear) map from the manifold into itself that enables one to move between different points on the manifold. Given images  $I_1(\mathbf{x})$  and  $I_2(\mathbf{x})$ , (where  $I(\mathbf{x})$  denotes the intensity at the spatial location  $\mathbf{x} = (x, y) \in [0, 1] \times [0, 1]$ ) a *transport operator*  $T$  is a mapping that acts in the following fashion:

$$I_2(\mathbf{x}) = I_1(\mathbf{x} + T(\mathbf{x})). \quad (1)$$

Instead of relying on the linear tangent space that accounts only for infinitesimal transformations, nonlinear transport operators can be well-defined over larger regions on the IAM.

For certain classes of articulations, the associated transport operators have algebraic structure in the form of a Lie group [11, 10, 12, 4]. In such instances, it is possible to explicitly compute transport operators that capture the curved geometric structure of an IAM. An example of this is the case of *affine* articulations, where the transport takes the form

$$I(\mathbf{x}) = I_0(A\mathbf{x} + \mathbf{t}).$$

In this case, the transport operator  $T(\cdot)$  is of the form  $T(\mathbf{x}) = (A - \mathbb{I})\mathbf{x} + \mathbf{t}$ ; this can be modeled as the group of 2D affine transformations. The affine group has found extensive use in tracking [12] and registration [14]. Miao and Rao [10] learn affine transport operators for image ensembles using a matrix exponential-based generative model and demonstrate improved performance over locally linear approximations. Culpepper and Olshausen [4] extend this framework using a more complex model on the transport operator in order to model paths on image manifolds.

Other common examples of articulations in computer vision are the projective group (used to model homographies and projective transformations) and diffeomorphisms (used to model 1D warping functions, density functions) [15]. How-

ever, while algebraic transport methods are mathematically elegant, they are applicable only to a very restrictive class of IAMs. Many IAMs of interest in computer vision and image processing applications, including IAMs corresponding to 3D pose articulations and non-rigid deformations, possess no explicit algebraic structure.

#### 1.4. Optical flow-based transport

In this paper, we study a specific class of transport operators that are generated by the *optical flow* between images (we introduced this notion empirically in [13]). Given two images  $I_1$  and  $I_2$ , the optical flow between them is defined to be the tuple  $(\mathbf{v}_x, \mathbf{v}_y) \in L^2([0, 1]^2) \times L^2([0, 1]^2)$  such that

$$I_2(x, y) = I_1(x + \mathbf{v}_x(x, y), y + \mathbf{v}_y(x, y)). \quad (2)$$

A common assumption in computing the optical flow between images is *brightness constancy* [16], where the spatial intensity is assumed to not change between  $I_1$  and  $I_2$ . Since the pioneering work of Horn and Schunk [16], there has been significant progress towards the robust estimation of optical flow between image pairs [17, 18].

Optical flow is a natural and powerful transport operator to transform one image into another. In the context of image manifolds, the collection of all optical flow operators at a point on an IAM is a manifold of the same dimension as the IAM [13]. In other words, at any reference image on an IAM, there is a corresponding manifold of optical flow operators that transports the reference image along the IAM. This new operator manifold, which we christen the *optical flow manifold* (OFM), can be used to obtain a canonical chart for the IAM [13]; this enables significantly improved navigation capabilities on the IAM than previous methods. In particular, OFM-based transport is well-defined even in instances when the transport operators cannot be modeled as a Lie group; an example of this is the pose manifold (the IAM associated with rigid body motion).

To see the efficacy of optical flow-based transport, consider again the case of image interpolation, but now with optical flow as the transport operator (see Fig. 3). Here, the path  $\gamma(t), t \in [0, 1]$  generated on the IAM via optical flow is a better representative of a path on the IAM, i.e.,  $\gamma(t) \in M$  for all  $t$ . Moreover, if the IAM is generated via Lie group actions, then this path coincides with the geodesic.

As discussed earlier, IAMs composed of images with sharp edges and textures lack smoothness and hence do not support locally linear modeling. In contrast, for a large class of interesting articulations, including affine transformations and

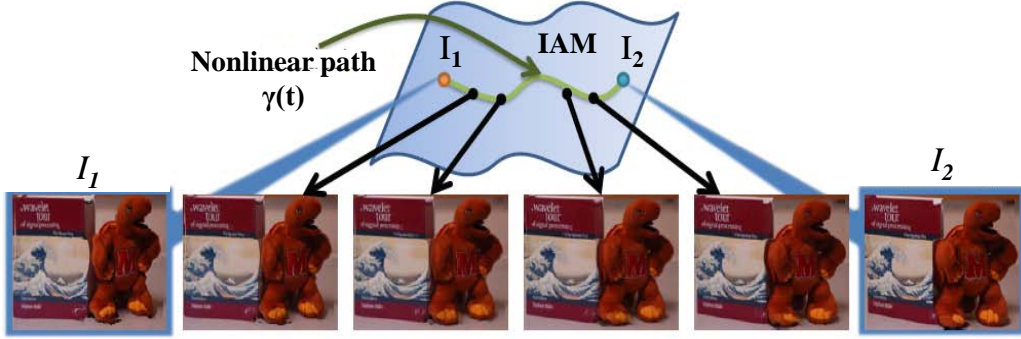


Figure 3: Optical flow based transport on an IAM leads to accurate image interpolation. Consider again the interpolation task in Fig. 1. While locally linear transport on IAMs lead to inaccurate interpolation results, the path  $\gamma(t)$  generated using optical flow provides an accurate approximation to the true path between  $I_1$  and  $I_2$ . Moreover, if the IAM is generated via Lie group actions, then the path generated coincides with the geodesic.

3D pose, the corresponding OFMs are smooth (see Appendix A and [13]) and support local linear modeling. Moreover, we can define the distance between two optical flows to measure the amount of motion required to articulate from one image to another. These properties suggest that machine learning techniques (such as LLE, ISOMAP, etc.) should be able to extract a considerable amount of geometrical information about an image ensemble when applied to its OFMs. Figure 4 confirms this fact by showcasing the improvement in OFM vs. IAM-based dimensionality reduction for an object under 2D translation.

The stability of OFM-based modeling enables us below to develop new differential geometric tools for signal processing such as the geometric mean and Karcher mean [19] of an image ensemble. The Karcher mean of a set of images is defined as the image on the IAM that is closest to the set in the sense of *geodesic distance*. Figure 5 shows the results of Karcher mean estimation using the OFM. The interested reader is referred to [13] for details of the Karcher mean estimation as well as more OFM-based tools and applications.

### 1.5. Related work

OFMs fall under the category of deformation modeling, which has been studied in many different contexts, including active shape [20] and active appearance models [21]. However, there are significant differences between OFMs and traditional deformation modeling approaches. Morphlets [22], for example, provide a multiscale modeling of and interpolation across image deformations, but their

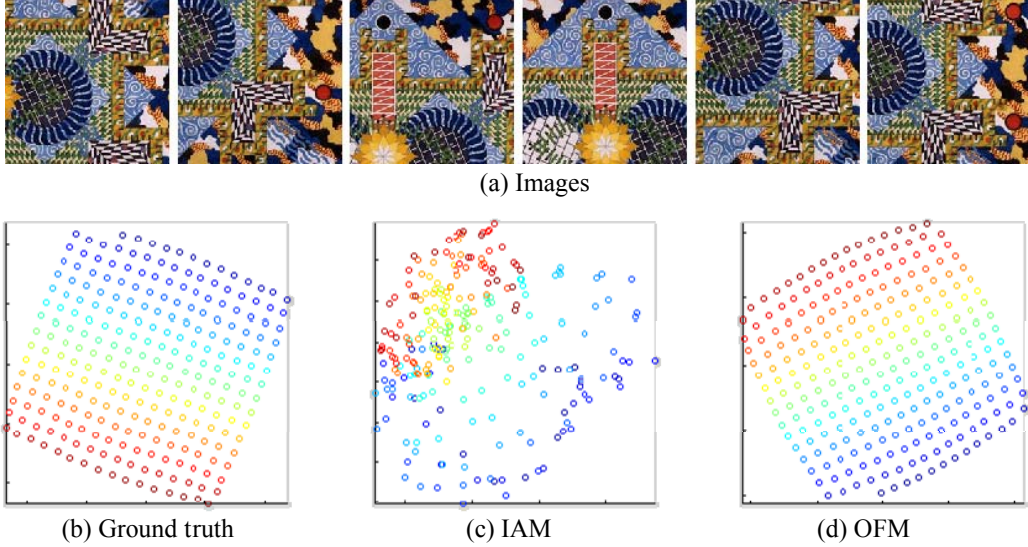


Figure 4: Comparison of manifold learning on a IAM vs. OFM (from [13]). The IAM is generated by cropping patches of size  $200 \times 200$  pixels at random from an image — thereby generating a 2D translation manifold. (a) Sample images from the IAM showing a few images at various translations. (b-d) 2D embedding obtained using ISOMAP on (b) the ground truth, (c) the IAM, and (d) the OFM. The near perfect embedding obtained using OFM hints at the near-perfect isometry in the OFM distances.

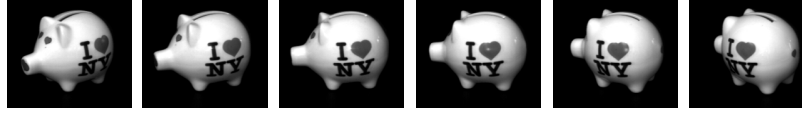
treatment is limited to image pairs. In contrast, OFMs apply to image ensembles consisting of a potentially large number of images. Beymer and Poggio [23] have argued for the use of motion-based representations for learning problems. However, their goal is image synthesis, and therefore they offer no insights into the geometric nature of manifold-valued data. Jojic et al. [24] use a layered representation to represent videos by separating the appearance of moving objects from their motion and then representing each using subspace and manifold models, respectively. This simple, yet powerful, representation can model and synthesize complex scenes using simple primitives, but it is not intended to go beyond simple manifolds such as those generated by translations and affine transformations.

### 1.6. Specific contributions

The main aim of this paper is to develop the mathematical foundation of optical flow-based transport operators for image manifolds, which were introduced empirically [13].

We first define a metric on an IAM using its corresponding OFMs. Each OFM





(a) Sampling of images from an IAM



(b) Ground truth



(c) IAM-based  
Karcher mean



(d) OFM-based  
Karcher mean

Figure 5: Karcher mean estimation for 20 images generated by the rotation of an object about a fixed axis [13]. The images are from the COIL dataset [25]. (a) Sample images from the IAM showing a few images at various rotations. We estimate the Karcher mean using local linear transport on the IAM and the OFM. Shown above are (b) the ground truth Karcher mean of the images, (c) the Karcher mean estimated using local linear transport on the IAM, and (d) the Karcher mean estimated using local linear transport on the OFM. The accuracy of the estimate obtained from the OFM showcases the validity of local linear transport on the OFM.

has a natural metric that is a locally isometric function of the corresponding parameter values. We consider the induced metric on the IAM, which we dub the *flow metric* and show that the flow metric between two points of the IAM is a measure of the distance between the corresponding parameter values. Next, using the flow metric, we develop analytic notions of curvature, optical flow fields, and parallel transport. We analyze in detail the case of optical flow fields defined along a fixed curve. In particular, we define a unique function associated to each such flow field, which we dub the *motion function*, using which we can define the notion of *parallel* flow fields. We answer the natural question of how one can optimally approximate a non-parallel field by a parallel field, and thereby induce uniform motion along the curve. We also construct a monoid structure on the set of all flow fields along a fixed curve. Under certain conditions on the curvature of a curve, we show that the space of parallel optical flow fields along the curve is a submonoid that comes with a convenient multi-scale structure.

We envision that the theory developed in this paper will enable a large class of practical applications involving image manifolds especially under sparse sampling of images from the manifold. In this context, we believe that the long term impact of this paper is the first step towards a complete theory of manifolds for arbitrary classes of signals using transport operators along the lines of the classic

differential geometry for smooth manifolds.

### 1.7. Organization

The remainder of the paper is organized as follows. In Section 2, we introduce OFMs and, using a fixed metric on the OFMs, study the induced flow metric on the corresponding IAM. We compare dimensionality reduction techniques on the IAM vs. OFMs. Using the flow metric, we develop geometric tools on the IAM in Section 3 and highlight their application to parameter estimation. We also develop error bounds on approximating non-parallel flow fields by parallel flow fields and illustrate the idea with the example of video resampling. In Section 4 we develop the multiscale structure of parallel flow fields. We conclude in Section 5 with a brief discussion.

## 2. OFMs and the Flow Metric

In this section, we define and study the basic properties of OFMs corresponding to an IAM  $M$ . Much of the section is concerned with a formal introduction to OFMs, first defined in [13], leading towards the development of our fundamental tool, the flow metric on an IAM.

### 2.1. Optical flow manifolds

The optical flow between two images on an IAM measures the apparent motion between the two images and thus reflects the corresponding parameter change between the two images. For a fixed base image  $m \in M$ , consider a neighborhood  $N(m)$  of  $m$ . If for  $m' \in N(m)$  there exist flow vectors  $(\mathbf{v}_x, \mathbf{v}_y)$  such that  $m'$  can be obtained from  $m$  using the flow vector, i.e.,  $m'(x, y) = m(x + \mathbf{v}_x, y + \mathbf{v}_y)$ , then we say that optical flow exists from  $m$  to  $m'$ . We denote this situation as  $\phi_{\mathbf{v}_x, \mathbf{v}_y}(m) = m'$  or simply  $\phi(m) = m'$ . In practice, occlusion or boundary effects (i.e., veiling of certain portions of an image due to changes in the background or interferers in the scene) may lead to undefined estimates for the flow vectors. However, as described in [13], one can mitigate these issues by incorporating additional consistency tests to ensure that only the meaningful flow vectors are retained.

The set of all points  $m' \in M$  for which optical flow from  $m \in M$  to  $m'$  exists is a neighborhood of  $m$ , which we denote by

$$B_m = \{m' \in M : m'(x, y) = m(x + \mathbf{v}_x, y + \mathbf{v}_y)\}.$$

Using this neighborhood, we define the optical flow manifold  $O_m$  at  $m$  as follows.

**Definition 2.1.** Let  $M$  be a  $K$ -dimensional IAM. Given  $m \in M$ , the Optical Flow Manifold  $O_m$  is defined as the set of optical flows  $\phi$  between  $m$  and points in  $B_m$

$$O_m = \{\phi = (\mathbf{v}_x, \mathbf{v}_y) : \phi(m) \in B_m\}.$$

The neighborhood  $B_m$  is called the flow neighborhood around  $m$ .

$O_m$  is nonlinear, i.e., an arbitrary OFM is not always a linear vector space. It is clear that the collection of neighborhoods  $\{B_m : m \in M\}$  covers  $M$ .

**Proposition 2.2.** Let  $M$  be a  $K$ -dimensional IAM. Then  $O_m$  has the structure of a manifold homeomorphic to  $B_m$ .

*Proof.* We first show that  $O_m$  is homeomorphic to  $B_m$  and pullback the smooth structure of  $B_m$  to induce a smooth structure on  $O_m$ . Consider the map  $g : B_m \rightarrow O_m$  that sends a point  $m' \in B_m$  to  $\phi \in O_m$  such that  $\phi(m') = m$ .  $g$  is clearly bijective and hence, we use  $g$  to endow  $O_m$  with the quotient topology. Moreover, since  $g$  is injective,  $g$  is a homeomorphism since  $g$  is an open map: given an open set  $U \subset B_m$  we have  $U = g^{-1}(g(U))$  and by the quotient topology on  $O_m$ , we have that  $V = g(U)$  is open in  $O_m$  since  $g^{-1}(V) = U$  is open in  $B_m$ . Since  $\{B_m : m \in M\}$  is a chart for the manifold structure of  $M$ , there exists continuous maps  $\{\Phi_m : m \in M\}$  so that  $\{(B_m, \Phi_m) : m \in M\}$  is an atlas for  $M$ . Therefore, the composition  $\Phi_m \circ g^{-1}$  is a homeomorphism from  $O_m$  to an open set  $W$  in  $\mathbb{R}^K$ . Since  $O_m$  is now covered by the single chart  $\{O_m, \Phi_m \circ g^{-1}\}$ , we can pullback this smooth to  $O_m$  via  $\Phi_m \circ g^{-1}$  and endow  $O_m$  with a smooth structure. Moreover,  $B_m$  is homeomorphic to  $O_m$  by construction.  $\square$

We make a few preliminary observations. First, by the homeomorphic relationship between  $O_m$  and  $B_m$ , we conclude that  $O_m$  is also a  $K$ -dimensional manifold. Also,  $O_{m_1}$  is homeomorphic to  $O_{m_2}$  for any  $m_1, m_2 \in M$ . Moreover, the trivial element  $\phi_0 \in O_m$  for each  $m \in M$  that maps  $m$  to itself in  $B_m$  acts a natural “origin” in  $O_m$ .

We observe, again from the homeomorphic relationship between  $B_m$  and  $O_m$ , that given any point  $m_1 \in B_m$  there is a unique operator  $\phi_1$  such that  $\phi_1(m) = m_1$ . A question to ask is whether there is any relationship between  $O_m$  and the tangent space  $T_m$  at  $m$ . Indeed, for a smooth IAM, one can define tangent spaces  $T_m$  at  $m \in M$  and the OFM  $O_m$  at  $m$  is diffeomorphic to a neighborhood of  $0 \in T_m$ . To see this fact, recall that the *exponential map* defined on  $T_m$  is a diffeomorphism between a neighborhood  $U_0$  of  $0 \in T_m$  and a neighborhood  $V_m$  of  $m$ . We also have a diffeomorphism  $\Phi$  from  $O_m$  to  $B_m$  by definition. Now,  $B_m$  is open and

there is an open ball  $B_0 \subseteq U_0$  on which  $\Phi^{-1} \circ \exp$  is a diffeomorphism, being a composition of two diffeomorphisms.

As a concrete example, consider again the translational manifold  $M_T$  from Section 1.2. Note that the parameter space in this case is  $\Theta = \mathbb{R}^2$ . Since there is no occlusion between any two images on  $M_T$ , it follows that given any  $m, m' \in M_T$ , there exists a  $\phi \in O_m$  such that  $\phi(m) = m'$ . Since this is true for any pair of images in  $M_T$ , we conclude that  $O_m = \mathbb{R}^2 \forall m \in M_T$  and hence,  $B_m = M_T \forall m \in M_T$ . More generally, we note that for an IAM  $M$  generated by Lie group actions without occlusion between images, the OFM  $O_m$  at any point  $m \in M$  can be identified with the parameter space  $\Theta$  and neighborhood  $B_m$  at each point is the entire manifold  $M$ . In particular, one can recover the geodesic path between two any two points  $m, m' \in M$  by using appropriate flow operators in  $O_m$  to generate the geodesic path from  $m$  to  $m'$ . This shortest path corresponds to the geodesic in the parameter space between the parameter values corresponding to  $m, m'$  as well. We will return to this example in future sections and show how our more general formulation contains the algebraic methods such as [11, 10, 12, 4] as special cases.

As in the case of the tangent bundle, we can construct an analogous bundle with the collection of  $O_m$  since  $m$  varies through  $M$ .

**Definition 2.3.** Let  $M$  be an IAM and  $O_m$  the OFM at  $m \in M$ . The flow bundle  $OM$  on  $M$  to be the disjoint union of the  $O_m$  since  $m$  varies over  $M$

$$OM = \coprod_{m \in M} O_m.$$

Thus, an element of  $OM$  is a pair  $(m, \phi)$  with  $\phi \in O_m$ . Using this fact, we can induce a topology on the flow bundle.

**Proposition 2.4.** Given a  $K$ -dimensional IAM  $M$ , the flow bundle  $OM$  is a  $2K$ -dimensional manifold.

*Proof.* We first note that if  $(m, \phi) \in OM$  then  $\{m, \phi(m)\} \in \{m\} \times B_m$ . Thus, given an atlas  $\{(U_\lambda, \psi_\lambda)\}_{\lambda \in \Lambda}$  of  $M$ , we have that

$$\psi_{\alpha, \beta}^*((m, \phi)) = (\psi_\alpha(m), \psi_\beta(\phi(m)))$$

maps  $(m, \phi)$  into  $\psi_\alpha(U_\alpha) \times \psi_\beta(U_\beta \cap B_m)$  where  $U_\alpha$  and  $U_\beta$  are charts around  $m$  and  $\phi(m)$  respectively. The (continuous) inverse of  $\{x_1, \dots, x_k, y_1, \dots, y_k\} \in \psi_\alpha(U_\alpha) \times \psi_\beta(U_\beta \cap B_m)$  is given by  $(m, \phi)$  where  $\psi_\alpha(m) = \{x_1, \dots, x_k\}$  and  $\phi$  is the unique operator in  $O_m$  such that  $\psi_\beta(\phi(m)) = \{y_1, \dots, y_k\}$  so that  $OM$  is locally Euclidean.  $\square$

Thus, we see that OFMs are manifolds consisting of flow operators that are defined pointwise on the corresponding IAM. The action of flow operators at a base point on the IAM results in motion *along* the IAM, as opposed to linear transport that results in motion *off* the manifold.

The key property that makes the study of OFMs interesting is that, for interesting IAMs, the associated OFMs are smooth and exhibit nice distance properties [13]. We summarize these in Appendix A. These two properties, namely smoothness and isometry, are in turn used to define a meaningful distance on the IAM. We discuss this next.

## 2.2. Metric structure on an IAM via its OFMs

Consider again the translational manifold  $M_T$ , where the OFM  $O_m$  at each point  $m \in M_T$  can be identified with  $\mathbb{R}^2$ . Being isometric with  $\mathbb{R}^2$ , we can endow each  $O_m$  with the Euclidean metric which we denote as  $d_O(\cdot, \cdot)$ . Let  $\theta_1, \theta_2 \in \mathbb{R}^2$  be a pair of parameters such that  $m_1 = i(\theta_1), m_2 = i(\theta_2)$ ; note that there exists a  $\phi \in O_{m_1}$  such that  $\phi(m_1) = m_2$ . It then follows that  $d_O(\phi_0, \phi) = C\|\theta_1 - \theta_2\|$  for some  $C > 0$ . As the results of [13] indicate, the above discussion holds analogously for generic OFMs, i.e., each  $O_m$  has an associated metric  $d_O(\cdot, \cdot)$  and this metric is locally isometric to a corresponding metric on the parameter space  $\Theta$ . We indicate this as

$$d_O(\phi_0, \phi) \propto d_\Theta(\theta_1, \theta_2),$$

where  $\phi_0$  is the unique operator in  $O_{m_1}$  such that  $m_1 = \phi_0(m_1)$ .

Our main focus in the remainder of this section is to define a corresponding metric for IAMs using the metric  $d_O(\cdot, \cdot)$  on  $O_m$ . The resulting metric on  $M$  inherits the property of being locally isometric to the changes in parameters. As a first step, we locally “push forward” the metric on  $O_m$  onto  $B_m$  as follows. For points  $m_1, m_2 \in M$  with  $m_2 \in B_{m_1}$ , we have a unique operator  $\phi_1$  such that  $m_2 = \phi_1(m_1)$  so that we can define the distance  $d_M(m_1, m_2)$  as the corresponding distance between  $\phi_0$  and  $\phi_1$

$$d_M(m_1, m_2) := d_O(\phi_0, \phi_1).$$

Moreover, if  $m_1 = f(\theta_1)$  and  $m_2 = f(\theta_2)$  for parameters  $\theta_1, \theta_2 \in \Theta$  then we have

$$d_M(m_1, m_2) \propto d_\Theta(\theta_1, \theta_2).$$

However, this definition does not readily extend to the case where  $m_1$  and  $m_2$  are not “optically related”, i.e.,  $m_2 \notin B_{m_1}$ . In this case, we first connect  $m_1$  and  $m_2$

by a path  $c$  such that  $c(0) = m_1, c(1) = m_2$ . We then partition the domain of  $c$  by a partition  $P = \{0 = t_0 < t_1 < \dots < t_n = 1\}$  such that the intermediate points along the path are optically related i.e.,  $c(t_i) \in B_{c(t_{i-1})}$ , where we assume that, for a fine enough partition, we can obtain such a nesting. We then define the distance along  $c$  to be

$$d(c, m_1, m_2) = \sup_P \sum_{i=0}^{n-1} d_M(c(t_i), c(t_{i+1})),$$

where the supremum is over all partitions of the path. By taking the infimum over all possible paths, we obtain a metric on  $M$ , i.e.,

$$d_M(m_1, m_2) = \inf_c \sum_{i=0}^{n-1} d_M(c(t_i), c(t_{i+1})). \quad (3)$$

In essence, the metric  $d_M(\cdot, \cdot)$  is similar to the Riemannian distance: we first define distance over a fixed curve and then take the infimum over all possible paths (see Fig. 6).

**Proposition 2.5.** *For an IAM  $M$ , the distance  $d_M(\cdot, \cdot)$  in (3) is a metric on  $M$ .*

*Proof.* Positivity of  $d_M(\cdot, \cdot)$  is clear, as is the fact that  $d_M(m, m) = 0$ . If  $m_1 \neq m_2$ , then for every path  $c$  between  $m_1$  and  $m_2$ , we have that  $\sum_{i=0}^{n-1} d_M(c(t_i), c(t_{i+1})) \neq 0$  and hence,  $d_M(m_1, m_2) \neq 0$ . Symmetry follows from the fact that along  $c$ ,  $c(t_i) \in B_{c(t_{i-1})}$  and hence,  $d_M(c(t_i), c(t_{i+1})) = d_M(c(t_{i+1}), c(t_i))$ . For the triangle inequality, we note that given paths  $c_1$  and  $c_2$  from  $m_1$  to  $m_2$  and  $m_2$  to  $m_3$  respectively, the path  $c_1 * c_2$  obtained by traversing  $c_1$  and  $c_2$  in succession at twice the rate (i.e.  $c_1 * c_2(t) = c_1(2t)$  for  $0 \leq t \leq \frac{1}{2}$  and  $c_1 * c_2(t) = c_2(2t - 1)$  for  $\frac{1}{2} \leq t \leq 1$ ) is a path from  $m_1$  to  $m_3$  and  $\sum_{i=0}^{n-1} d_M(c_1(t_i), c_1(t_{i+1})) + \sum_{i=0}^{n-1} d_M(c_2(t_i), c_2(t_{i+1})) \geq \sum_{i=0}^{n-1} d_M(c_1 * c_2(t_i), c_1 * c_2(t_{i+1}))$ . By taking the infimum over all such paths, we verify the triangle inequality.  $\square$

Recall that the metric on  $O_m$  satisfies  $d_O(\phi_0, \phi) \propto d_\Theta(\theta_1, \theta_2)$  with  $\phi \in O_m$  and  $\theta_1, \theta_2$  the parameters corresponding to  $m$  and  $\phi(m)$  respectively. From the above result, we see that, with the metric  $d_M(\cdot, \cdot)$  on  $M$ , we have

$$d_M(m_1, m_2) \propto d_\Theta(\theta_1, \theta_2),$$

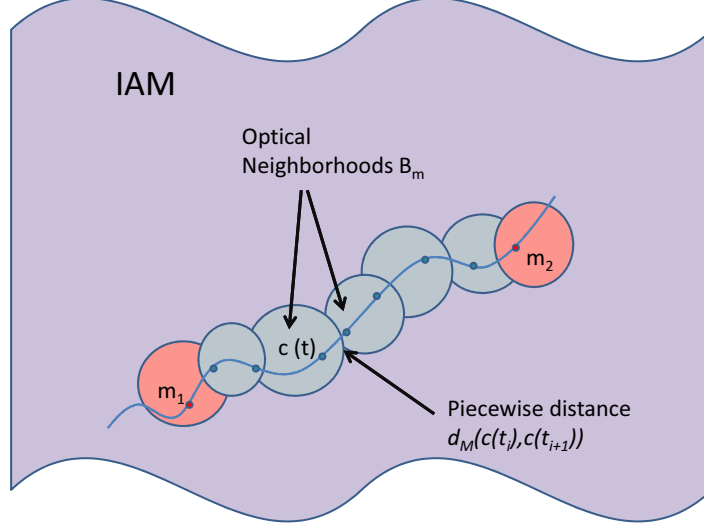


Figure 6: A pictorial representation of the OFMs and the flow metric for an IAM. Each OFM is mapped homeomorphically onto the flow neighborhood  $B_m$  of the corresponding base point. The optical flow distance between two points is computed by considering the infimum over all possible curves of the piecewise flow distance  $d_M(c(t_i), c(t_{i+1}))$ .

where  $m_1$  and  $m_2$  are points on  $M$  corresponding to the parameter values  $\theta_1$  and  $\theta_2$  respectively under the assumption that the monotonicity of  $d_O(\cdot, \cdot)$  over different  $\{O_m\}$  is universal, i.e., the distance between two points  $m_1$  and  $m_2$  does not change when we change our definitions of flow neighborhoods. We refer to this metric on  $M$  as the *flow metric*. We note that the flow metric is dependent on the OFMs  $O_m$ . We can now state the main result of this section.

**Theorem 2.6.** *Let  $M$  be an IAM and let  $d_M(\cdot, \cdot)$  be the associated flow metric. Then,  $d_M(m_1, m_2) \propto d_\Theta(\theta_1, \theta_2)$  as defined in (3) with  $m_1$  and  $m_2$  points on  $M$  corresponding to the parameter values  $\theta_1$  and  $\theta_2$  respectively.*

The smoothness of the OFMs makes them amenable for use with conventional manifold processing tools. For instance, common dimensionality reduction methods such as [5, 6, 7], etc. assume that the manifold is smooth and, hence, are tools that are more appropriate for use with the flow metric. As an example, consider Fig. 4. Here, we see that the pairwise flow metric between points varies smoothly, as opposed to the Euclidean metric on the IAM, where due to sharp edges, the matrix of pairwise distance has large off-diagonal entries. Moreover, the residual error of non-linear dimensionality reduction using the flow metric decays very

rapidly. As a result, it is more tractable to analyze IAMs using the flow metric as opposed to the conventional Euclidean metric.

As an example, consider the problem of estimating  $\theta \in \Theta$  such that  $i(\theta) = m$  given a finite number of template points  $m_1, \dots, m_n \in M$  with neighborhoods  $B_{m_1}, \dots, B_{m_n}$  that cover  $M$ . As a motivating special case, consider first the situation when  $M = M_T$ , the translational manifold with  $\Theta = \mathbb{R}^2$  and a single template point  $m'$ . A similar problem has been dealt with in [10, 4] where the authors estimate the Lie group generators corresponding to the IAM. Given the parameter value  $\theta' = (c'_1, c'_2)$  of the base image  $m'$ , we can compute the parameter  $\theta$  corresponding to  $m$  as follows. First, we find the unique flow operator  $\phi = (\phi_1, \phi_2) \in \mathbb{R}^2$  such that  $\phi(m') = m$ . The parameter value  $\theta$  corresponding to  $m$  is then obtained as  $\theta = (c'_1 + \phi_1, c'_2 + \phi_2)$ .

An entirely similar result holds for generic IAMs generated by Lie group actions. From this simple example, we see that finding the optimal  $\theta$  is equivalent to finding the optimal flow operator  $\phi$  that minimizes  $d_M(\phi(m'), m)$  with  $\phi \in O_{m'}$ .

We return to the general case with  $n$  template images and seek the optimal flow operator  $\phi \in O_{m_i}$  that minimizes  $d_M(\phi(m_i), m)$ ,  $i = 1, \dots, n$ . Here, a single neighborhood does not cover the entire IAM, and, hence, to estimate  $\phi$  we first find the neighborhood  $B_{\tilde{m}} \in \{B_{m_1}, \dots, B_{m_n}\}$  such that  $\tilde{m} = \arg \min_{m' \in \{m_1, \dots, m_n\}} d_M(m, m')$ . Our search is then restricted to the neighborhood  $B_{\tilde{m}}$ . Within this neighborhood, we find the optimal  $\phi \in O_{\tilde{m}}$  as above, i.e.,  $\phi = \arg \min_{\phi \in O_{\tilde{m}}} d_M(\phi(\tilde{m}), m)$ . In essence, we first find the optimal template point and then search within the corresponding OFM for the optimal flow operator. If a single template point generates the entire IAM, then this procedure clearly reduces to the Lie group case discussed earlier. We thus see that the generic OFM formulation includes the algebraic methods of [10, 12] as a special case.

The remainder of the paper is devoted to developing geometric tools for IAMs that leverage the flow metric. Note that, unlike the tangent space, the OFM has no linear structure and, hence, we do not have at our immediate disposal tools such as parallel translation, covariant derivatives etc. We will construct analogous tools for our purposes via the flow metric and hence open up a vista for IAM analysis.

### 3. Geometric Tools for IAMs via the Flow Metric

In this section, we develop the basic tools needed to analyze the structure of IAMs using the flow metric. We will pay special attention to flow operators defined along curves on the corresponding IAMs. Keeping computations in mind,



these tools will open the door to a variety of applications as mentioned in the Introduction and beyond.

### 3.1. Flow radius

We first seek an appropriate measure of the size of an OFM and the corresponding flow neighborhood. In classical differential geometry, one measures the radius at a point in terms of the injectivity radius [9] using the Riemannian metric. In a similar fashion, we will measure the radius of a point  $m$  in an IAM  $M$  in terms of the flow metric.

#### 3.1.1. Flow Radius at $m \in M$

**Definition 3.1.** Given an IAM  $M$  and  $m \in M$  define the flow radius or simply the radius  $r_m$  at  $m$  as

$$r_m = \sup_{n \in B_m} d_M(m, n).$$

Note that we may regard  $r_m$  as a function from  $M$  to  $\mathbb{R}^+$  i.e.  $r(m) = r_m$  is a map from  $M$  into the non-negative reals. Moreover, it is continuous as a function of  $m$ . Consider the variations of  $r_m$  as  $m$  varies. If  $r_m$  is large, then one can find a suitable operator  $\phi \in O_m$  that transports  $m$  to a far away point, with distance measured using the flow metric. Conversely, a small  $r_m$  indicates that  $m$  can only be transported within a small region, or said differently,  $m$  obstructs transport on the IAM. Moreover, rapid changes in the magnitude of  $r_m$  within a small neighborhood indicate that the manifold is not well-behaved near  $m$ . In particular, this indicates that there are several points close to  $m$  that obstruct transport while there are also several points that allow flow over large distances on the IAM.

#### 3.1.2. Flow radius for Lie groups

A class of IAMs for which  $r_m$  is very well behaved, indeed for which  $r_m$  is a constant, are those generated by a Lie group action. As a motivating example, consider again the translational manifold  $M_T$ . We note that as  $O_m = \mathbb{R}^2$  for each  $m \in M_T$ , it follows that  $r_m = \infty$ . If the parameter space  $\Theta$  is compact, for instance, if  $\Theta = S^1$  and we consider affine rotations of a base image  $m$  generating the IAM  $M$ , then we again have that  $B_m = M$  and  $r_m$  is a constant  $0 < c < \infty$ . For a generic IAM  $M$  generated by Lie group actions with  $B_m = M$ , it follows that the flow radius is a constant whose exact value depends both on the object being imaged as well as the nature of the articulation.

### 3.1.3. Flow curvature

The above discussion indicates that the reciprocal  $\frac{1}{r_m}$  is a measure of “curvature”. When  $r_m$  is large (or infinite), then the IAM is can be thought as being “optically flat” at  $m$  in the sense that there is no obstruction to transport on the manifold at  $m$ . In the other limiting case, i.e., if  $r_m$  approaches zero, the only operator in  $O_m$  is the trivial operator  $\phi_0$  and hence,  $M$  is has high “curvature” at  $m$ . We thus have the following definition.

**Definition 3.2.** *Given an IAM  $M$  and  $m \in M$  define the flow curvature or simply the curvature at  $m$  as  $K_m = \frac{1}{r_m}$ .*

While the traditional notion of curvature is a point property, the flow curvature depends on both the base point as well as its neighborhood properties. In terms of flow curvature, we can now state that if  $M$  is generated by a Lie group then  $M$  has constant curvature. The class of IAMs with constant curvature will play a prominent role in our later analysis.

### 3.2. Optical flow fields

In this section, we focus on a construct motivated by differential geometry, namely the idea of a vector field on a manifold. Recall that a vector field is a *section* of the the tangent bundle, i.e., a vector field is a map  $\sigma : M \rightarrow TM$  such that  $\pi \circ \sigma = id_M$ , where  $\pi$  is the natural projection from the tangent bundle and  $id_M$  is the identity map on  $M$ . In an analogous fashion, we define an *optical flow field*, or simply a *flow field*, as a section of the *optical flow bundle* of  $M$ .

While vector fields are defined generically on manifolds, the special class of vector fields along curves is especially important in differential geometry. Vector fields along curves give rise to tools such as parallel translation, Jacobi fields, etc. [9]. In the case of IAMs, which lack analytic structure in general, we will define optical fields along curves on an IAM and recover similar geometric tools using the flow metric. As we assign operators to points on the curve, we intuitively would like the transport induced by the operator to remain along the curve so that the collection of flow operators induces motion along the curve. This special class of optical fields will be our main object of study for the rest of the paper.

**Definition 3.3.** *Let  $M$  be an IAM and let  $c$  be a smooth curve passing through  $m_1, m_2 \in M$ . Define  $B_{c(t)} \cap c = \{c(t') \in O_{c(t)} : t' \geq t\}$ . An optical field from  $m_1$  to  $m_2$  along  $c$  with  $m_1, m_2 \notin \partial c$  is a map  $V : t \mapsto OM$  such that  $V(t) := V_t \in O_{c(t)}$  and  $V_t(c(t)) \in \overline{B_{c(t)} \cap c}$  where by  $V_t(c(t))$  we mean the point on the curve obtained by the action of  $V_t$  on  $c(t)$ .*

In other words, an optical field along  $c$  is an assignment of a flow operator  $V_t$  with  $V_t$  an element of the OFM at  $c(t)$  such that the action  $V_t$  on  $c(t)$  remains on the curve. Thus, the action of  $V_t$  on the base point  $c(t)$  at time  $t$  induces motion along the curve. When  $m_1$  and  $m_2$  are clear from the context, we simply refer to  $V_t$  as the optical field along  $c$ .

In essence, the curve  $V_t$  traces a curve in  $OM$  as  $t$  varies with a consistent action on  $c(t)$ . By  $B_{c(t)} \cap c$  we mean the intersection of the  $K$ -dimensional flow neighborhood  $B_{c(t)}$  with the one dimensional curve  $c$  starting at  $c(t)$  and hence,  $B_{c(t)} \cap c$  is a one dimensional embedded curve in  $B_{c(t)}$  i.e.,  $B_{c(t)} \cap c$  is a one dimensional “slice” of  $B_{c(t)}$ . We assume that  $B_{c(t)} \cap c$  is connected.

To measure the distance traveled by the action of  $V_t$  on  $c(t)$ , we define a radius  $r_t$  restricted along the curve as opposed to the complete flow radius  $r_{c(t)}$  at  $c(t)$

$$r_t = \sup_{n \in B_{c(t)} \cap c} d_c(c(t), n),$$

where  $d_c(\cdot, \cdot)$  is the flow metric restricted to the curve  $c$ . Likewise, the curvature  $K_t$  along  $c$  is the ratio

$$K_t = \frac{1}{r_t}.$$

Since  $B_{c(t)} \cap c$  is connected and 1D, the distance  $d_c(c(t), m)$  between  $c(t)$  and  $m \in B_{c(t)} \cap c$  characterizes  $m$  in the following sense. Given any positive constant  $0 \leq \eta \leq r_t$  there is a unique  $m \in B_{c(t)} \cap c$  with  $d_c(c(t), m) = \eta$ . Thus, by specifying the distance along the curve  $c$ , we effectively characterize the curve. Lifting this observation into  $O_{c(t)}$ , we have the following result.

**Theorem 3.4.** *Let  $M$  be an IAM, let  $c$  be a curve passing through  $m_1, m_2 \in M$ , and let  $V_t$  be an optical field from  $m_1$  to  $m_2$  along  $c$ . Then,  $V_t$  is completely characterized by the function  $h_V(t) = d_c(V_t(c(t)), c(t))$  in the sense that for any non-negative function  $h(t)$  bounded pointwise by  $r_t$ , there exists a unique optical field  $V_t$  such that  $h_V(t) = h(t)$ .*

*Proof.* That  $0 \leq h_V(t) \leq r_t$  is clear from the definitions above. Let  $h(t)$  be any non-negative function bounded by  $r_t$ . Then, for fixed  $t = t_0$  we have that  $0 \leq h(t_0) < r_{t_0}$ . By the remark made previously, we have a unique  $m_{t_0} \in B_{c(t_0)} \cap c$  with  $d_c(c(t_0), m_{t_0}) = h(t_0)$ . Now, as  $m_{t_0} \in B_{c(t_0)} \cap c$ , in particular,  $m_{t_0} \in B_{c(t_0)}$  and hence there exists a unique  $\phi_{t_0} \in O_{c(t_0)}$  such that

$$\phi_{t_0}(c(t_0)) = m_{t_0}.$$

As  $t_0$  was arbitrary, as  $t$  varies, we can define  $V_t = \phi_t$ . Moreover,

$$h_V(t) = d_c(V_t(c(t)), c(t)) = d_c(\phi_t(c(t)), c(t)) = d_c(m_t, c(t)) = h(t)$$

so that  $h(t)$  characterizes  $V_t$ . □

Recall that we view an optical flow field  $V_t$  along  $c$  as a curve in  $OM$ . Theorem 10 states that this curve is characterized by the function  $h_V(t)$  in the sense that for a given  $V_t$  along the (fixed) curve  $c$ , the function  $h_V(t)$  contains all the information about the motion induced by  $V_t$  on the curve i.e.,  $h_V(t)$  measures the distance to which  $V_t$  transports  $c(t)$  in  $B_{c(t)} \cap c$ . Since this function is of prime importance, we make the following definition.

**Definition 3.5.** *Given an optical field  $V_t$  along a curve  $c$ , define its motion function to be*

$$h_V(t) = d_c(V_t(c(t)), c(t)).$$

Note that by continuity of the metric  $d_c(\cdot, \cdot)$  and  $c(t)$ , the function  $h_V(t)$  is also continuous.

### 3.3. Parallel flow fields

There is a very natural geometric interpretation of the motion function  $h_V$  of an optical flow field  $V_t$  along a curve  $c$ . Namely, it is a measure of the distance traveled along the curve at time  $t$  by  $c(t)$  when acted upon by  $V_t$ . Thus, it is natural to think of instantaneous changes in  $h_V$  in  $t$  as a measure of the *velocity* of the motion induced by  $V_t$  on  $c(t)$ . In classical geometry, the class of constant velocity curves is especially important; they correspond to uniform motions. Similarly, we will be interested in the class of constant motion functions  $h_V$ . These correspond to optical fields  $V_t$  along  $c$  that induce uniform motion along  $c$ , where by uniform motion we mean the distance traveled along  $c$  is constant for all time.

Another key link with classical differential geometry is the notion of parallel transport (or parallel translation) of tangent vectors [9]. Parallel transport is a key tool for “moving” tangent vectors from different tangent spaces while preserving direction/orientation along the curve. Parallel transport utilizes the inherent linear structure of the tangent space to define a linear map between tangent on points along the curve. We aim to develop a similar analytic tool for IAMs. However, an immediate stumbling block is the clear lack of *linear structure* in the OFM. We therefore take a different approach to defining parallel translation in the OFM case using motion functions.

To motivate our definition, we recall that a vector field along a curve  $c$  on a manifold is *parallel* if its *covariant derivative* along  $c$  vanishes. The covariant derivative is in essence a way of differentiating the vector field along  $c$ . The relation between parallel translation and parallel vector fields is that, given a tangent vector  $v$  in the tangent space of a point  $c(t)$  on the curve  $c$ , it is possible to extend  $v$  along  $c$  by parallel translation to yield a parallel vector field along  $c$ . In the OFM case, we have the motion function of an optical field along  $c$  at our disposal and we use it to characterize parallelism of the field along the curve.

**Definition 3.6.** *An optical field  $V_t$  along a curve  $c$ , is parallel if the derivative of its motion function with respect to  $t$  is zero, or equivalently, if the motion function is constant along  $c$ , i.e.,  $h_V(t)$  is a constant.*

In all that follows, we will denote by  $\Omega(c, m_1, m_2)$  the space of all optical fields through  $m_1$  and  $m_2$  along a curve  $c$  passing through  $m_1, m_2 \in M$ . When  $m_1, m_2$  are understood from context, we simply refer to this space as  $\Omega(c)$ . The subclass of parallel fields along  $c$  will be denoted by  $\omega(c)$ .

A few facts are immediate from the above definition. First, since  $h_V(t) \leq r_t$  for all  $t$ , it is clear that if  $V_t$  is to be parallel along  $c$ , then  $h_V(t)$  is a constant  $h_V$  independent of  $t$  and  $h_V \leq \inf_t r_t$ . Therefore, in the future, we will suppress the argument  $t$  in the motion function  $h_V(t)$  of a parallel flow field. Second, since  $V_t$  is characterized by  $h_V$ , we see that for any constant  $\delta$  such that  $0 \leq \delta < \inf_t r_t$ , there is a parallel optical field along  $c(t)$  such that  $h_V = \delta$ . Such a parallel optical field can be obtained, for example, by choosing for each  $t$  a flow operator  $\phi_t \in O_{c(t)}$  with  $d_c(\phi_t(c(t)), c(t)) = \delta$ . The existence of such an optical field  $\phi_t$  is guaranteed by the above theorem. Given  $\phi \in O_{c_0}$  we aim to extend  $\phi$  throughout the curve to obtain a flow field  $V_t$  such that  $V_t$  is parallel along  $c$  with  $h_V = d_c(\phi(c_0), (c_0))$ . The following result allows for such parallel translation of flow operators along a curve.

**Proposition 3.7.** *Let  $M$  be an IAM, and let  $c$  be a curve through  $m_1, m_2 \in M$ . Let  $\epsilon = \inf_t r_t$ . Then, given any  $\phi \in O_{c(t_0)}$  with  $\delta = d_c(\phi(c(t_0)), c(t_0)) < \epsilon$  there exists a unique parallel optical field along  $c(t)$  with  $h_V(t) = \delta$ .*

*Proof.* As  $\delta < \epsilon$ , in particular  $\delta < r_t$ . Thus, invoking the previous theorem for the special case of the constant function  $h(t) = \delta$ , we have the existence of a unique optical field  $V_t$  along  $c(t)$  such that  $h_V(t) = \delta$ . Moreover, since  $h_V(t)$  is constant,  $V_t$  is parallel.  $\square$

In contrast with the classical case, parallel transport along a curve is dependent on the nature of the flow operator, i.e., an arbitrary flow operator  $\phi \in O_{c_0}$  cannot be parallel translated along  $c$  unless  $d_c(\phi(c_0), c_0) < \inf_t r_t$ . This constraint is related to the nature of the curve; parallel transport along a curve that contains points with high curvature  $K_t$  is limited to those flow operators that induce smaller motion along the curve. Moreover, the possibility of parallel transport of  $\phi \in O_{c(t)}$  has a global dependence, i.e., it depends on the curvature of the entire curve, not only the curvature  $K_t$  at the point  $c(t)$ .

Thus, we are naturally led to study those curves for which parallel translation of an operator  $\phi$  at a single point  $c(t_0)$  ensures the existence of parallel translation of operators at any other point on the curve. Clearly, the necessary condition is the invariance of  $K_t$  with  $t$ , and therefore we are led to consider curves for which the curvature  $K_t$  is independent of  $t$ , i.e., a constant. This special class of curves has a very rich structure that we explore in the following sections.

### 3.4. Approximation of arbitrary flow fields by parallel flow fields

In this section, we consider the problem of approximating an arbitrary  $V_t \in \Omega(c)$  by elements in  $\omega(c)$ . Consider an optical field  $V_t \in \Omega(c)$  (not necessarily parallel) along a curve  $c$  of constant flow curvature. A natural question to ask is how far away  $V_t$  is from being parallel. One way to do this is to seek the “best” approximation of  $V_t$  by a parallel field  $W_t \in \omega(c)$  along  $c$ . To quantify the approximation, we consider the following cost

$$e(t) = d_c(V_t(c(t)), W_t(c(t))).$$

In words,  $e(t)$  is a measure of how the action of  $V_t$  on  $c(t)$  differs from the action of  $W_t$  on  $c(t)$ . Note that  $e(t)$  is bounded above by  $h_V(t) + h_W$  since  $e(t) = d_c(V_t(c(t)), W_t(c(t))) \leq d_c(V_t(c(t)), c(t)) + d_c(c(t), W_t(c(t))) = h_V(t) + h_W$ .

While  $e(t)$  is a pointwise error, we will need to consider the total error over the entire curve. To this end, a natural choice of error metric is

$$E(V, W) = \int_a^b e(t) dt,$$

where the domain of  $c$  is the interval  $(a, b)$ . Thus, our goal is to find a parallel optical field  $W_t \in \omega(c)$  that minimizes  $E(V, W)$ , i.e.,

$$W^* = \arg \min_W E(V, W).$$

In general, a minimizer may not exist, or it may not be unique if one exists. However, the greatest issue is the strong dependence of the error  $e(t)$  on the flow metric, which prevents a generic solution to the minimization problem since we cannot infer the convexity of the problem as stated. We can, however, obtain a universal lower bound on the error independent of the flow metric as follows

$$|h_V(t) - h_W| = |d_c(V_t(c(t)), c(t)) - d_c(W(c(t)), c(t))| \leq e(t).$$

Therefore, we seek a minimizer of  $\int_a^b |h_V(t) - h_W| dt$ . Note that since  $W_t$  is parallel,  $h_W$  is a constant, say  $h \geq 0$ . Moreover, since  $c$  is a path of constant flow curvature,  $r_t$  is a constant  $r > 0$ . Since  $h$  characterizes  $W_t$ , a minimizer  $h^*$  yields a lower bound for  $E(V, W)$ . Our goal then, is to find an optimal constant  $h^*$  that minimizes  $\widetilde{E(h)} = \int_a^b |h_V(t) - h| dt$ . We claim that a solution  $h^*$  is a certain “median” of  $h_V(t)$ . We first define  $A_k = \{t \in (a, b) : h_V(t) > k\}$  and  $B_k = \{t \in (a, b) : h_V(t) \leq k\}$  for some  $k$ . We claim that an optimal constant  $\hat{h}$  is such that  $\lambda(A_{\hat{h}}) = \lambda(B_{\hat{h}})$ , where  $\lambda(S)$  denotes the measure of a set  $S$ .

**Theorem 3.8.** *Let  $\hat{h}$  be the constant such that  $\lambda(A_{\hat{h}}) = \lambda(B_{\hat{h}})$ . Then,  $\hat{h}$  minimizes  $\widetilde{E(h)}$ .*

*Proof.* Note that the function  $\widetilde{E(h)}$  is convex in  $h$  with  $h \in (a, b)$  and  $(a, b)$  a convex interval. Therefore, we are guaranteed a minimizer  $h^*$ . Now, without loss of generality, we assume that  $h^* < \hat{h}$ . We evaluate the cost over the two regions  $A_{\hat{h}}$  and  $B_{\hat{h}}$

$$\widetilde{E(h^*)} = \int_a^b |h_V(t) - h^*| dt = \int_{A_{\hat{h}}} |h_V(t) - h^*| dt + \int_{B_{\hat{h}}} |h_V(t) - h^*| dt.$$

Define  $B_1 = \{t : h^* < h_V(t) < \hat{h}\}$  and  $B_2 = \{t : h_V(t) < h^*\}$  and note that  $\lambda(A_{\hat{h}}) - \lambda(B_1) - \lambda(B_2) = 0$ . We can now express  $\widetilde{E(h^*)}$  in terms of  $\widetilde{E(\hat{h})}$  as

$$\widetilde{E(h^*)} = \widetilde{E_{A_{\hat{h}}}(\hat{h})} + (\hat{h} - h^*)\lambda(A_{\hat{h}}) + \widetilde{E_{B_{\hat{h}}}(\hat{h})} - (\hat{h} - h^*)\lambda(B_2) - \alpha,$$

where  $\widetilde{E_X(\hat{h})}$  denotes the cost function restricted to the subset  $X \subseteq (a, b)$  and  $\alpha$  is a positive constant that measures the difference of  $\hat{h} - h^*$  on the set  $B_1$ . The maximum value of  $\alpha$  is  $(\hat{h} - h^*)\lambda(B_1)$  and hence

$$\widetilde{E(h^*)} \leq \widetilde{E_{A_{\hat{h}}}(\hat{h})} + (\hat{h} - h^*)(\lambda(A_{\hat{h}}) - \lambda(B_1) - \lambda(B_2)) + \widetilde{E_{B_{\hat{h}}}(\hat{h})}.$$

Since  $\lambda(A_{\hat{h}}) - \lambda(B_1) - \lambda(B_2) = 0$ , we conclude that

$$\widetilde{E(h^*)} \leq \widetilde{E(\hat{h})}$$

and hence  $h^* = \hat{h}$ . The case of  $h^* > \hat{h}$  follows from symmetry.  $\square$

We illustrate the approximation of non-parallel flow fields by parallel flow fields with the example of *video resampling* where we can consider a video  $\mathcal{I}$  to be a curve on an IAM i.e., a video  $\mathcal{I} = \{I_t, 0 \leq t \leq T\}$  with  $T > 0$ . This application is related to the problem of dynamic time warping (DTW) [26, 27, 28], where one is interested in measuring the similarity between two sequences that vary in time or speed. As we shall see, this can be used for matching or aligning video sequences with a warped time axis [27, 28].

Consider the IAM generated by imaging a black disk on an infinite white background starting with an initial velocity  $v_0$  and accelerating with constant acceleration  $a$  along a fixed direction. For instance, the disk can be thought of as undergoing freefall off an infinitely high cliff. The IAM is a 1D curve  $c$  and homeomorphic to  $\mathbb{R}^+$ , the non-negative reals. Note that the curvature is everywhere zero since  $r_t = \infty$ . Given an arbitrary flow field  $V_t$  along  $c$ , our goal is to analytically construct a parallel field  $\tilde{V}_t$  such that  $\tilde{V}_t$  is the unique parallel flow field that minimizes  $\widetilde{E(h)}$ . We first consider the video obtained by the action of  $V_t$  i.e.  $\mathcal{I} = \{V_t(c(t))\}$ . Since  $V_t$  is not parallel, the video will show the disk moving with non-uniform motion. Our goal is to make the video uniform, i.e., generate a new video  $\tilde{\mathcal{I}}$  from  $\tilde{V}_t$  that shows the disk moving with uniform motion.

From the physics of the problem, it is clear that  $d_c(c(t), c(t + \delta_t)) = K(v_t\delta_t + \frac{1}{2}a\delta_t^2)$  for some positive constant  $K$  and any time increment  $\delta_t$  with  $v_t$  being the velocity at time  $t$ . Now, since  $\tilde{V}_t$  is to be a parallel flow field, we have that  $h_{\tilde{V}}$  is a constant denoted by  $h$ . Thus,  $d_c(\tilde{V}_t(c(t)), c(t)) = d_c(c(t), c(t + \delta_t)) = h$  so that  $K(v_t\delta_t + \frac{1}{2}a\delta_t^2) = h$ . Rearranging this equation, we arrive at

$$\delta_t^2 + \frac{2v_t}{a}\delta_t - \frac{2h}{aK} = 0.$$

Solving for  $\delta_t$ , we obtain two real roots  $\delta_t^{1,2} = \frac{-v_t}{a} \pm \sqrt{(\frac{v_t}{a})^2 + \frac{2h}{aK}}$  of which the (physically meaningful) positive root  $\delta_t^1 = \frac{-v_t}{a} + \sqrt{(\frac{v_t}{a})^2 + \frac{2h}{aK}} > 0$  is retained. Thus, by defining  $\tilde{V}_t$  such that  $\tilde{V}_t(c(t)) = c(t + \delta_t^1)$ , we see that  $\tilde{V}_t$  is the unique parallel flow field that minimizes  $\widetilde{E(h)}$ . The new video  $\tilde{\mathcal{I}} = \{\tilde{V}_t(c(t))\}$  will thus show the disk moving with constant velocity. With this, we have effectively linearized the motion and made it independent of the acceleration of the disk.



## 4. Multiscale Structure of Parallel Flow Fields

As indicated in the Section 3, the set of parallel optical fields is a very special subset of the set of all optical fields along a fixed curve  $c$ . In this section, we will construct a monoid structure (i.e., a set with an associative operation and identity) on the set of all optical fields along a fixed curve  $c$  and show that the class of parallel fields forms a submonoid of this set under some conditions on the curvature along the curve. Moreover, the monoid operation yields a multiscale structure on the set of parallel optical fields.

### 4.1. Monoid structure on $\Omega(c)$

As noted previously in Section 2, a clear disadvantage in dealing with the space of all optical fields along a curve is the lack of a linear, or more generally, any algebraic structure. In order to remedy this situation, we will define a binary operation on the set of optical fields that yields a monoid structure. We first fix a curve  $c$  passing through  $m_1, m_2 \in M$ .

Recall that a generic  $V_t \in \Omega(c)$  is characterized by its motion function  $h_V(t)$ . Thus, operations defined on motion functions  $h_V(t)$  translate to operations on  $V_t \in \Omega(c)$ . With this in mind, we define for  $V_t, W_t \in \Omega(c)$  the sum  $V_t + W_t$  to be the unique optical field with motion function

$$h_{V+W}(t) = \min(h_V(t) + h_W(t), r_t).$$

Since  $h_{V+W}(t) \leq r_t$  for all  $t$ , we see that  $h_{V+W}(t)$  corresponds to a unique flow field that we define to be  $V_t + W_t$ . This operation is clearly commutative. Note also that the trivial (parallel) field  $Z_t$  defined to be the field that acts trivially on  $c(t)$ ; i.e.,

$$Z_t(c(t)) = c(t)$$

is characterized by the motion function  $h_Z(t) = 0$  since  $h_Z(t) = d_c(Z_t(c(t)), c(t)) = d_c(c(t), c(t)) = 0$ . Moreover, for any  $V_t \in \Omega(c)$ , we have that

$$Z_t + V_t = V_t.$$

We see that  $Z_t$  acts as the identity element in  $\Omega(c)$ . In addition to this, there is also the unique optical field  $U_t$  characterized by

$$h_U(t) = r_t$$

that satisfies

$$V_t + U_t = U_t$$

for all  $V_t \in \Omega(c)$  and hence acts as the “absorbing” element of  $\Omega(c)$ .

Clearly, we do not have “inverses” with respect to “+” in the sense that, given a generic  $V_t \in \Omega(c)$ , there does not exist a  $W_t$  such that  $V_t + W_t = Z_t$ . However, we do have “conjugates” with respect to “+” in the following sense. Given any  $V_t \in \Omega(c)$ , there is a unique  $V_t^* \in \Omega(c)$  such that

$$V_t + V_t^* = U_t.$$

$V_t^*$  is defined by its motion function

$$h_{V^*}(t) = r_t - h_V(t).$$

We refer to  $V_t^*$  as the conjugate of  $V_t$ . Moreover,

$$(V_t^*)^* = V_t.$$

Finally, we note that  $U_t$  and  $Z_t$  are conjugates.

**Proposition 4.1.**  $\Omega(c)$  is a monoid under the operation “+” defined in (22).

*Proof.* We only verify associativity, since  $Z_t$  provides the identity. Given  $V_t, X_t, Y_t \in \Omega(c)$ , we consider the sums  $(V_t + X_t) + Y_t$  and  $V_t + (X_t + Y_t)$ . If  $h_V(t), h_X(t), h_Y(t)$  are such that  $h_V(t) + h_X(t) + h_Y(t) < r_t$  then both sides are clearly equal. If on the other hand,  $h_V(t) + h_X(t) + h_Y(t) \geq r_t$  then we must have that either the sum  $h_V(t) + h_X(t) + h_Y(t)$  taken two factors at a time exceeds  $r_t$  or that the combined sum of all three factors exceeds  $r_t$  with the sum of no two factors exceeding  $r_t$ . In the first case, we assume that  $h_V(t) + h_X(t) \geq r_t$ , which implies the sum

$$(V_t + X_t) + Y_t = U_t$$

with  $(X_t + Y_t) \neq U_t$ , i.e.,

$$h_X(t) + h_Y(t) < r_t.$$

However,  $V_t + (X_t + Y_t)$  is characterized by

$$\min(h_V(t) + \min(h_X(t) + h_Y(t), r_t), r_t) = \min(h_V(t) + h_X(t) + h_Y(t), r_t) = r_t$$

which shows that

$$V_t + (X_t + Y_t) = U_t$$

as well. The other cases follow by similar arguments.  $\square$

#### 4.2. Parallel fields along curves of constant curvature

The addition operation defined above restricts to the set  $\omega(c)$  of parallel fields along  $c(t)$ . However, for generic  $c(t)$ , the sum  $V_t + W_t$  of two parallel fields  $V_t, W_t \in \omega(c)$  may result in a possibly non-parallel field. For instance, consider  $h_V(t), h_W(t)$  such that  $h_V(t_0) + h_W(t_0) < r_{t_0}$  but  $h_V(t_1) + h_W(t_1) \geq r_{t_1}$  for some  $t_0, t_1$ . Clearly, the sum is not parallel since in the first case,  $h_{V+W}(t_0) = h_V(t_0) + h_W(t_0) < r_{t_0}$  while  $h_{V+W}(t_1) = r_{t_1}$  so that  $h_{V+W}(t)$  is not constant.

However, that if  $r_t$  is constant along  $c$ , i.e., the flow radius is constant along the path, then the above situation is vacuous. Since  $K_t = \frac{1}{r_t}$  is the curvature along the curve  $c$ , we are essentially requiring the curvature to be constant along  $c$ . We formally record the above observation.

**Proposition 4.2.** *Let  $c$  be a curve with constant curvature  $K_t$ . Then, the operation “+” restricted to  $\omega(c)$  is well defined and  $\omega(c)$  is a submonoid of  $\Omega(c)$ .*

*Proof.* We need only verify closure of “+” in  $\omega(c)$  since  $Z_t \in \omega(c)$ . If  $V_t \in \omega(c)$ , then  $h_V(t)$  is a constant and we will therefore suppress the argument  $t$  in  $h_V(t)$ . Note that since  $c$  is of constant curvature,  $r_t$  is a constant  $r > 0$ . Thus, if  $V_t, W_t \in \omega(c)$ , then  $h_{V+W} = \min(h_V + h_W, r)$ . If  $h_V + h_W < r$  then as  $h_V$  and  $h_W$  are both constant so is their sum  $h_V + h_W$  and

$$h_{V+W} = h_V + h_W.$$

If  $h_V + h_W \geq r$ , then

$$h_{V+W} = \min(h_V + h_W, r) = r.$$

In either case,  $h_{V+W}$  is constant and hence

$$V_t + W_t \in \omega(c).$$

□

Thus, curves  $c$  that have constant curvature at all points are very special; not only is parallel transport determined by a single point on the curve, but  $\omega(c)$  is also a submonoid of  $\Omega(c)$ .

#### 4.3. Multiscale structure of $\omega(c)$

In this section, we look for finite submonoids of  $\omega(c)$ . We would like the submonoids to be canonically defined, by which we mean that they depend only on the geometry of the constant curvature curve  $c$ . In particular, we construct for each positive integer  $k$ , a collection of finite submonoids  $\mathcal{V}_{n,k}$  of  $\omega(c)$  with  $n$  a non-negative integer and  $c$  a constant curvature path. Moreover, we will see that this collection is naturally nested, i.e.,

$$\mathcal{V}_{0,k} \subset \mathcal{V}_{1,k} \subset \dots$$

In essence, this provides a multi-scale view of  $\omega(c)$ .

Since  $\omega(c)$  consists of parallel optical fields, we will simply denote the motion function  $h_V(t)$  of  $V_t \in \omega(c)$  as  $h_V$ . Moreover, we will suppress the subscript  $t$  when referring to elements of  $V_t \in \omega(c)$ . Also, for  $V \in \omega(c)$ , we mean by  $\frac{V}{k}$  the element in  $\omega(c)$  with motion function being the constant  $\frac{h_V}{k}$ . Recall also that  $\omega(c)$  possesses two canonical elements  $Z, U \in \omega(c) \cap \Omega(c)$  that act as the trivial and absorbing elements of  $\omega(c)$  respectively.

We begin now with the construction of  $\mathcal{V}_{n,k}$ . Fix a positive integer  $k$  and set

$$\mathcal{V}_{0,k} = \{Z, U\}.$$

Next, we inductively set

$$\mathcal{V}_{n,k} = \left\{ Z, \frac{U}{k^n}, \frac{2U}{k^n}, \dots, \frac{(k^n-1)U}{k^n}, U \right\}.$$

Now, it is clear that

$$\mathcal{V}_{0,k} \subset \mathcal{V}_{1,k} \subset \dots$$

Moreover, by the constant curvature condition, the sum of any two elements in  $\mathcal{V}_{n,k}$  remains in  $\mathcal{V}_{n,k}$  while associativity is obtained from the corresponding property in  $\omega(c)$ . Finally, since  $Z \in \mathcal{V}_{n,k}$  for all  $n, k$ , we conclude that each  $\mathcal{V}_{n,k}$  is a finite submonoid of  $\omega(c)$ .

We have thus obtained a sequence of finite submonoids of  $\omega(c)$ . Moreover, with increasing  $n$ , it is clear that an arbitrary  $V \in \omega(c)$  can be uniformly approximated by  $\tilde{V} \in \mathcal{V}_{n,k}$  in the sense that  $|h_V - h_{\tilde{V}}|$  can be made arbitrarily small by choosing larger  $n$ . In other words, the sequence  $\mathcal{V}_{n,k}$  is “dense” in  $\omega(c)$ . Finally, since the basic generators of  $\mathcal{V}_{n,k}$  are  $Z, U \in \omega(c) \cap \Omega(c)$ , this construction is canonical in the sense that it depends only on the set of  $\{Z, U\}$ , which are in turn are characterized by the global geometry of the curve  $c$ .

In Section 3, we saw that with curvature conditions on a curve  $c$ , it is possible to approximate an arbitrary flow field by a parallel one. The results presented above show that one can further approximate a parallel flow field with a finite collection of “template” elements from  $\omega(c)$ . This is very much similar to the multiscale representation that wavelets provide for natural images. Indeed, the multiscale structure inherent to monoids paves the way for “lossy compression” of arbitrary flow fields by storing only the relevant scales  $\mathcal{V}_{n,k}$ . In addition to compression, the multiscale structure can potentially enable fast computations on the flow fields. Operations on the flow fields can equivalently be mapped to those on the  $\mathcal{V}_{n,k}$  without loss in accuracy while gaining significantly in the number of computations required.

## 5. Discussion

In this paper, we have developed the mathematical foundations of optical flow-based transport operators for image manifolds, which were introduced empirically [13]. Our main theoretical contribution was the development of the flow metric for using the ambient metric on the OFMs. Using the flow metric, we derived differential geometric analogues of tangent bundles, vector fields, parallel transport, curvature etc. When the IAM is generated by Lie group parameters, we showed that the OFM framework includes previous algebraic methods as a special case. Moreover, since the flow neighborhood at each point is the entire IAM, we can obtain geodesics between any two images using flow operators. While this paper has focused on explaining and extending the results obtained in [13], we envision that the theory could make an on a large class of applications involving image manifolds.

A clear assumption in our analysis has been that optical flow is the transport operator of choice for IAMs. While this is true for a majority of IAMs generated by motion-induced parameter changes such as translations, rotations and unstructured plastic deformations, there are classes of IAMs for which the choice of transport operator is not immediately clear. For instance, for illumination manifolds obtained by variations in the illumination of an object, optical flow may not be the transport operator of choice, since such manifolds do not in general obey the brightness constancy requirement needed in optical flow computations. Moreover, in cases where there is significant self-occlusion during the imaging process, optical flow may not be a practical transport operator. However, by regularizing the optical flow computation to handle occlusions by removing flow operators that lead to undefined motion between pixels, one can partially circumvent this

issue, and the methods of this paper can be profitably applied. However, these are aspects of a more computational nature, and we reserve them for an alternate forum.

A number of avenues for future work exist. First, our work hints that it should be possible to develop additional geometric tools such as affine connections, holonomy, etc. using flow operators. These will complete the IAM analysis toolbox. Second, although our development has been specific to the case of IAMs with optical flow, the basic model is extensible to a wide variety of signal ensembles with appropriately defined transport operators. For instance, a manifold model for speech signals has been proposed in [29] where appropriate transport operators and the analog of OFMs may involve a frequency domain approach. Once the “right” transport operator has been identified for the application in hand, one can conceivably define and study metrics similar to our flow metric and thereby develop analytic tools for further analysis.

## Acknowledgments

The authors thank Chinmay Hegde for valuable discussions and insightful comments on the manuscript.

This work was partially supported by the grants NSF CCF-0431150, CCF-0728867, CCF-0926127, CCF-1117939, ARO MURI W911NF-09-1-0383, W911NF-07-1-0185, DARPA N66001-11-1-4090, N66001-11-C-4092, N66001-08-1-2065, AFOSR FA9550-09-1-0432, and LLNL B593154.

## Appendix A. Smoothness and Distance Properties of OFMs

In this appendix, we establish the smoothness and isometry properties of OFMs associated with two interesting classes of IAMs: the affine articulation manifold and the pose manifold.

### Appendix A.1. Affine articulation manifold

Affine articulations are parameterized by a 6-dimensional articulation space  $\Theta = \mathbb{R}^6$ ; each articulation  $\theta$  can be written as  $\theta = (A, \mathbf{t})$ , with  $A \in \mathbb{R}^{2 \times 2}$  and  $\mathbf{t} \in \mathbb{R}^2$ . Any image belonging to the IAM can thus be written as

$$I_1(\mathbf{x}) = I_{\text{ref}}((A + \mathbb{I})\mathbf{x} + \mathbf{t}), \quad (\text{A.1})$$

where  $\mathbf{x} = (x, y)$  is defined over the domain  $\mathcal{X} = [0, 1] \times [0, 1]$  and where  $I_{\text{ref}}$  is a reference image.<sup>2</sup> The optical flow field  $f_\theta$  associated with the transport operator can be written as  $f_\theta(\mathbf{x}) = f_{A, \mathbf{t}}(\mathbf{x}) = A\mathbf{x} + \mathbf{t}$ . Now, recall that the OFM at  $I_{\text{ref}}$  is defined as

$$O_{I_{\text{ref}}} = \{f_\theta : \theta = (A, \mathbf{t}), A \in \mathbb{R}^{2 \times 2}, \mathbf{t} \in \mathbb{R}^2\}.$$

The linear dependence of the optical flow field on both  $A$  and  $\mathbf{t}$  implies that the OFM, which is the collection of optical flows at  $I_{\text{ref}}$ , is infinitely smooth. Finally, noting that the OFM is independent of the reference image, we can establish the following result.

**Lemma 1.** *For affine articulations, the OFM at any reference image is infinitely smooth.*

Next, we consider distances between optical flows and establish that the OFM is isometric.

**Lemma 2.** *For affine articulations, the OFM at any reference image is globally isometric.*

*Proof.* Given two articulations  $\theta_1 = (A_1, \mathbf{t}_1)$  and  $\theta_2 = (A_2, \mathbf{t}_2)$ , the Euclidean distance between them is given by

$$\begin{aligned} (d(f_{\theta_1}, f_{\theta_2}))^2 &= \int_{\mathbf{x} \in \mathcal{X}} \|f_{\theta_1}(\mathbf{x}) - f_{\theta_2}(\mathbf{x})\|_2^2 d\mathbf{x} \\ &= \int_{\mathbf{x} \in \mathcal{X}} \|(A_1 - A_2)\mathbf{x} + (\mathbf{t}_1 - \mathbf{t}_2)\|_2^2 d\mathbf{x} \\ &= (\theta_1 - \theta_2)^T \Sigma (\theta_1 - \theta_2), \end{aligned}$$

where

$$\Sigma = \int_{\mathbf{x} \in \mathcal{X}} \left( \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}^T & 1 \end{bmatrix} \right) d\mathbf{x}.$$

For  $\mathcal{X} = [0, 1] \times [0, 1]$ , it is easily shown that  $\Sigma$  is full-rank and positive definite. Hence, we have

$$d(f_{\theta_1}, f_{\theta_2}) = \|\theta_1 - \theta_2\|_\Sigma,$$

where  $\|\cdot\|_\Sigma$  is the Mahalanobis (or weighted Euclidean) distance defined using the matrix  $\Sigma$ . This implies that the OFM is Euclidean and, hence, globally isometric.  $\square$

---

<sup>2</sup>Boundary-related issues are always an concern when we define images over a finite domain. Here, we circumvent this by assuming that the regions of interest are surrounded by a field of zeros. This allows us to assign undefined values to zero and satisfy (A.1)

## Appendix A.2. Pose manifold

The pose manifold is the IAM corresponding to the motion of a camera observing a static scene. It is well-known that the articulation space is 6-dimensional, with 3 degrees of rotation and 3 degrees of translation of the camera, i.e.,  $\Theta = SO(3) \times \mathbb{R}^3$ . We assume that the optical flow is the (unique) 2-dimensional projection of the motion flow of the scene induced due to the motion of the camera.

Without loss of generality, we assume the reference articulation  $\theta_{\text{ref}} = (R_0, \mathbf{t}_0) = (\mathbb{I}, \mathbf{0})$  and that the camera's internal calibration is known and accounted for [30]. At the reference articulation  $\theta_{\text{ref}}$ , let the depth at a pixel  $\mathbf{x}$  be given by  $\lambda_{\text{ref}}(\mathbf{x})$ . Under an articulation  $\theta = (R_\theta, \mathbf{t}_\theta)$ , the optical flow observed at the pixel  $\mathbf{x}$  is given by

$$f(\theta)(\mathbf{x}) = P \left( \lambda_{\text{ref}}(\mathbf{x}) R_\theta \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} + \mathbf{t} \right) - P \left( \lambda_{\text{ref}}(\mathbf{x}) \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \right) \quad (\text{A.2})$$

where  $P(\cdot)$  is a projection operator such that  $P(x, y, z) = (x/z, y/z)$ . Note that  $P$  is a well-defined and *infinitely smooth function* provided  $z \neq 0$ .

The OFM  $\mathcal{O}$  is the image of the map  $f(\theta) = \{f(\theta)(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}$ , where  $\mathcal{X}$  is the image plane. We now establish the smoothness of the OFM.

**Lemma 3.** *Consider the OFM corresponding to a pose manifold of a scene with depth map (at the reference articulation  $\theta_{\text{ref}}$ ) strictly bounded away from zero, i.e.,  $\forall \mathbf{x}, \lambda_{\text{ref}}(\mathbf{x}) > \lambda > 0$ . Then there exists a neighborhood of  $\theta_{\text{ref}}$  where  $f(\theta) = \{f(\theta)(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}$  is an infinitely smooth map.*

*Proof.* The smoothness of  $f(\theta)$  follows from the smoothness of  $f(\theta)(\mathbf{x})$ , which in turn follows from the smoothness of the projection operator  $P$ . Note that  $P(x, y, z) = (x/z, y/z)$  is well defined and infinitely smooth only if  $z \neq 0$ . When all points in the scene have depth bounded away from zero, then there is guaranteed a neighborhood of  $\Theta$  around  $\theta_{\text{ref}}$  wherein all points continue to have depth bounded away from zero. This ensures that the projection of all points is well defined, and hence,  $f(\theta)$  is infinitely smooth in this neighborhood.  $\square$

Note that, in contrast to affine articulations, where the corresponding OFM is smooth globally, the OFM associated with the pose manifold is smooth only over a neighborhood of the reference point.

Next, we consider distances between optical flows and establish that the OFM is locally isometric. We begin by defining the following representation for rotation



matrices. Let  $\omega = (\omega_x, \omega_y, \omega_z) \in \mathbb{R}^3$ , and let  $\Omega_\omega \in \mathbb{R}^{3 \times 3}$  be the skew-symmetric matrix defined as

$$\Omega_\omega = \begin{bmatrix} 0 & -\omega_x & \omega_y \\ \omega_x & 0 & -\omega_z \\ -\omega_y & \omega_z & 0 \end{bmatrix}.$$

Noting that the matrix exponential  $e^{\Omega_\omega}$  is a rotation matrix, we define our articulations as  $\theta = (\omega, \mathbf{t}) \in \mathbb{R}^6$ , with  $\mathbf{t} = (t_x, t_y, t_z)$ .

Before we state a formal result, we need to introduce two assumptions that will help in simplifying the derivation. First, we assume that the imaging model is well-approximated by a weak perspective model [30]. In the weak perspective model, it is assumed that the variations in the scene depth are significantly smaller than the average depth of the scene. In such a case, the projection map  $P(\mathbf{x}) = (x/z, y/z)$  can be well-approximated by

$$P_{\text{wp}}(\mathbf{x}) = (x/z_{\text{av}}, y/z_{\text{av}}),$$

where  $z_{\text{av}}$  is the average scene depth. As a consequence, translations along the  $z$ -axis are unobservable [30]; hence, we restrict the articulation space to the rotation of the camera and translation along  $x$  and  $y$  axes alone. Second, we assume that the rotation undergone by the camera is small. This enables us to approximate the matrix-exponential  $e^{\Omega_\omega} \approx \mathbb{I} + \Omega_\omega$ . As a consequence of this assumption, we only obtain a local isometry result. With these two assumptions, we are ready to state a result on the distances between optical flows.

**Lemma 4.** *Consider the pose manifold under the assumption of the weak perspective imaging model. Then the OFM is locally isometric.*

*Proof.* Without loss of generality, we denote our reference articulation as  $\theta_{\text{ref}} = \mathbf{0}$ , which corresponds to identity rotation matrix and null translation vector. We are interested in the distance between two optical flow fields  $f(\theta_1)$  and  $f(\theta_2)$ . Under our assumptions of weak perspective imaging model and small rotation, we can compute the distance between the optical flows to be

$$\begin{aligned} & f(\theta_1)(\mathbf{x}) - f(\theta_2)(\mathbf{x}) \\ &= P_{\text{wp}} \left( \lambda_{\text{ref}} R_1 \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} + \mathbf{t}_1 \right) - P_{\text{wp}} \left( \lambda_{\text{ref}} R_2 \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} + \mathbf{t}_2 \right) \\ &= \frac{1}{z_{\text{av}}} \begin{bmatrix} (-\lambda_{\text{ref}}(\omega_{1,x}y - \omega_{1,y}) + t_{1,x}) - (-\lambda_{\text{ref}}(\omega_{2,x}y - \omega_{2,y}) + t_{1,x}) \\ (\lambda_{\text{ref}}(\omega_{1,x}x - \omega_{1,z}) + t_{1,y}) - (\lambda_{\text{ref}}(\omega_{2,x}x - \omega_{2,z}) + t_{2,y}) \end{bmatrix} \\ &= \frac{1}{z_{\text{av}}} \begin{bmatrix} -\lambda_{\text{ref}}(\omega_{1,x} - \omega_{2,x})y + \lambda_{\text{ref}}(\omega_{1,y} - \omega_{2,y}) + (t_{1,x} - t_{2,x}) \\ \lambda_{\text{ref}}(\omega_{1,x} - \omega_{2,x})x - \lambda_{\text{ref}}(\omega_{1,z} - \omega_{2,z}) + (t_{1,y} - t_{2,y}) \end{bmatrix} \end{aligned}$$

Notice that both  $t_{1,z}$  and  $t_{2,z}$  are absent in the expression above. As noted above, this is due to the assumption of weak perspective imaging model which makes the translation of the camera along the  $z$ -axis unobservable. A key observation is that the articulation parameters  $\theta_1 = (\omega_1, \mathbf{t}_1)$  and  $\theta_2 = (\omega_2, \mathbf{t}_2)$  can be expressed as a function that is linear in the expression above. Hence,

$$f(\theta_1)(\mathbf{x}) - f(\theta_2)(\mathbf{x}) = A(\mathbf{x}) \begin{bmatrix} \omega_1 - \omega_2 \\ \mathbf{t}_1 - \mathbf{t}_2 \end{bmatrix}.$$

Finally, summing over  $\mathbf{x} \in \mathcal{X}$ , we obtain

$$d(f(\theta_1), f(\theta_2))^2 = (\theta_1 - \theta_2)^T \left( \int_{\mathbf{x} \in \mathcal{X}} A^T(\mathbf{x}) A(\mathbf{x}) d\mathbf{x} \right) (\theta_1 - \theta_2).$$

Hence, the OFM is locally isometric.  $\square$

## References

- [1] D. L. Donoho, C. C. Grimes, Image manifolds which are isometric to euclidean space, *J. of Math. Imaging and Vision* 23 (2005) 5–24.
- [2] M. B. Wakin, D. L. Donoho, H. Choi, R. G. Baraniuk, The multiscale structure of non-differentiable image manifolds, in: *Proc. SPIE Wavelets XI*, 2005.
- [3] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, F. Warner, S. Zucker, Geometric diffusions as a tool for harmonic analysis and structure definition of data: *Diffusion maps* 102 (21) (2005) 7426–7431.
- [4] B. J. Culpepper, B. A. Olshausen, Learning transport operators for image manifolds, in: *Adv. in Neural Info. Proc. Sys.*, 2009.
- [5] J. Tenenbaum, V. de Silva, J. Landford, A global geometric framework for nonlinear dimensionality reduction, *Science* 290 (5500) (2000) 2319–2323.
- [6] S. Roweis, L. Saul, Nonlinear dimensionality reduction by locally linear embedding, *Science* 290 (5500) (2000) 2323–2326.
- [7] M. Belkin, P. Niyogi, Laplacian eigenmaps and spectral techniques for embedding and clustering, in: *Adv. in Neural Info. Proc. Sys.*, 2001.

- [8] Y. Ma, S. Soatto, J. Kosecka, S. Sastry, *An Invitation to 3-D Vision: From Images to Geometric Models*, Springer-Verlag, 2005.
- [9] M. P. do Carmo, *Riemannian Geometry*, Birkhäuser, 1992.
- [10] X. Miao, R. P. N. Rao, Learning the lie groups of visual invariance, *Neural computation* 19 (10) (2007) 2665–2693.
- [11] D. Lin, E. Grimson, J. Fisher, Learning visual flows: a lie algebraic approach, 2009.
- [12] O. Tuzel, F. Porikli, P. Meer, Learning on lie groups for invariant detection and tracking, in: *IEEE Conf. Comp. Vision and Patt. Recog.*, 2008.
- [13] A. C. Sankaranarayanan, C. Hegde, S. Nagaraj, R. G. Baraniuk, Go with the flow: Optical flow-based transport for image manifolds, in: *Allerton Conf. on Comm., Control and Computing*, 2011.
- [14] R. Li, R. Chellappa, Aligning spatio-temporal signals on a special manifold, in: *Euro. Conf. on Comp. Vision*, 2010.
- [15] M. I. Miller, L. Younes, Group actions, homeomorphisms, and matching: a general framework, *Intl. J. of Comp. Vision* 41 (1–2) (2001) 61–84.
- [16] B. K. P. Horn, B. G. Schunck, Determining optical flow, *Artificial intelligence* 17 (1-3) (1981) 185–203.
- [17] D. Sun, S. Roth, M. J. Black, Secrets of optical flow estimation and their principles, in: *IEEE Conf. Comp. Vision and Patt. Recog.*, 2010.
- [18] T. Brox, J. Malik, Large displacement optical flow: descriptor matching in variational motion estimation, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (3) (2011) 500–513.
- [19] H. Karcher, Riemannian center of mass and mollifier smoothing, *Comm. on Pure and Appl. Math.* 30 (5) (1977) 509–541.
- [20] T. F. Cootes, C. J. Taylor, D. H. Cooper, J. Graham, Active shape models-their training and application, *Comp. vision and image understanding* 61 (1) (1995) 38–59.

- [21] T. F. Cootes, G. J. Edwards, C. J. Taylor, Active appearance models, *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (6) (2001) 681–685.
- [22] J. Kaplan, D. Donoho, The morphlet transform: A multiscale representation for diffeomorphisms, in: *Workshop on Image Registration in Deformable Environments*, 2006.
- [23] D. Beymer, T. Poggio, Image representations for visual learning, *Science* 272 (5270) (1996) 1905–1909.
- [24] N. Jojic, B. J. Frey, Learning flexible sprites in video layers, in: *IEEE Conf. Comp. Vision and Patt. Recog.*, 2001.
- [25] S. A. Nene, S. K. Nayar, H. Murase, Columbia object image library (COIL-100), T.R. No. CUCS-006-96, Dept. Comp. Science, Columbia University.
- [26] M. E. Munich, P. Perona, Continuous dynamic time warping for translation-invariant curve alignment with applications to signature verification, in: *Intl. Conf. on Comp. Vision*, 1999.
- [27] K. Wang, T. Gasser, Alignment of curves by dynamic time warping, *Annals of Statistics* 25 (3) (1997) 1251–1276.
- [28] A. Bosco, A. Bruna, S. Battiato, G. Puglisi, Digital video stabilization through curve warping techniques, *IEEE Trans. on Consumer Electronics* 54 (2) (2008) 220–224.
- [29] A. Jansen, P. Niyogi, Intrinsic fourier analysis on the manifold of speech sounds, in: *ICASSP*, 2006.
- [30] R. Hartley, A. Zisserman, *Multiple view geometry*, Cambridge university press, 2000.